METHODS AND EXPERIMENTS IN MENTAL TESTS

C. A. RICHARDSON

STORAGE-ITEM
MAIN

LP9-L26G
U.B.C. LIBRARY

METHODS AND EXPERIMENTS IN MENTAL TESTS

33151 Clark. Univ. 28 Mar 31 1.75

Digitized by the Internet Archive
in 2010 with funding from
University of British Columbia Library

http://www.archive.org/details/methodsexperime00rich

METHODS AND EXPERIMENTS IN MENTAL TESTS

BY

C. A. RICHARDSON M.A. (Cantab.)

AUTHOR OF
"THE SIMPLEX GROUP INTELLIGENCE SCALE"

GEORGE G. HARRAP & CO. LTD.
LONDON CALCUTTA SYDNEY

First published May 1922
by GEORGE G. HARRAP & CO. LTD.
2 & 3 Portsmouth Street, Kingsway, London, W.C 2

Printed in Great Britain by Jarrold & Sons, Ltd., Norwich

PREFACE

THIS book is an attempt to provide, in a manner readily comprehensible to all those interested in the subject, an answer to some of the more important questions that have been asked as to the nature, validity, and methods of application of mental tests, and the conclusions to be drawn therefrom. It is not an attempt to theorize, but to illustrate general principles on a strictly empirical and experimental basis, and it is hoped that all who follow developments in education and in psychology may find some points of interest in it.

The experiment described in the *second* part of Chapter V was carried out in the county of Northumberland.

The note at the end of Chapter IV was published in *The British Journal of Psychology* (April 1922), and I am grateful to the Editor for permission to include it in this book.

Most of the chapters of the book were kindly read by Mr Cyril Burt, to whom I am much indebted for valuable comments.

C. A. R.

NEWCASTLE-ON-TYNE
April 1922

CONTENTS

CHAPTER		PAGE
I.	Introductory	9
II.	The Reliability of the Stanford-Binet Scale of Intelligence Tests as an Index of Educable Capacity	24
III.	The Derivation of Mental Ages from Scores in a Group Test	40
IV.	Methods of Estimating the True Intelligence Quotients of Adults and Adolescents	56
V.	The Reliability of the Group Intelligence Test as an Index of Educability	79
VI.	Conclusion	89
	Index	93

Methods and Experiments in Mental Tests

CHAPTER I

INTRODUCTORY

MENTAL tests are now becoming so familiar in the English educational world that it is unnecessary, in introducing a book such as this, to enter into a lengthy description of their nature. Preliminary experimental work has now, however, reached a stage where we may well pause to look round and to take stock of the results so far reached.

In the first place, it is necessary to consider briefly the criticisms which have been levelled at the tests. These criticisms fall into three categories according as they are directed at (1) supposed faults inherent in the tests themselves; (2) difficulties in carrying out the alleged aim of the tests; or (3) difficulty in defining what this aim really is.

The first kind of critic will glance through a test booklet, seize on some particular test, and say confidently, " This is far too difficult," or " This is far too easy," or " The mere sight of this will frighten the children (especially the nervous ones) out of their wits," or " How can you expect children to understand long words like this ? "

The answer to such a critic is simple and conclusive. He must be made to understand that the value of a test is not estimated by its appearance, but by what it is actually found to effect in practice. That is, he must be asked to look, not at the tests, but at their results. The procedure of mental testing is rigidly empirical. All tests, before any conclusions are drawn from them, are tried on a number of children sufficient to determine whether they are suitable or not. In the case of a test devised in this way criticisms of the first type are therefore met in advance.

Secondly, objections are sometimes urged against the possibility of fulfilling one primary aim of the tests, namely, to give a fair field and

no favour. This objection usually takes the form of an assertion that the tests can easily be ' coached ' for. It is true that in the present *individual* scales (such as the Binet scale) a few (not many) of the tests are markedly susceptible to the influence of coaching, but even here all who have had experience in giving such tests will know that the presence of coaching can be detected with but little difficulty. At most it is unlikely to add more than a few points to the child's score. Moreover, in *group* testing, which is the most practicable way of carrying out tests on a sufficiently large scale, the possible variety, not only of the actual items of the different tests, but also of the particular forms which those tests may take (this being in effect limited only by the inventiveness of the human mind), is so great that coaching is rendered a waste of time and labour. Even in a standardized group test which is placed upon the market the number and variety of items is sufficiently large to nullify the effects of coaching. In fact, the only point which can be sustained in this criticism is that, as children become familiar with mental

tests, the norms of performance may *slightly* improve owing to greater speed of getting to work and to absence of emotional disturbance due to confrontation with an unfamiliar kind of task. But any such change in the norms will, of course, be automatically accounted for in the interpretation of results.

Thirdly, it is frequently urged that although these tests are said to be tests of intelligence we are unable to define what intelligence is. Now, in the first place, it cannot be too strongly emphasized that our ability to define intelligence is a matter quite irrelevant to the question of the value of the tests. What is wanted is a method of separating the sheep from the goats —that is, the bright children, those who are capable of responding to our efforts to educate them, from the dull children, those whose capacity for response to educative influences is distinctly limited. If the tests do in fact enable us to grade children in order of their educability it does not matter in the least what we call them; though, since we commonly speak of bright children as 'intelli-

gent,' ' intelligence tests ' is probably as good a name as any.

But apart from this we can frame a fairly precise definition of intelligence without much difficulty. In fact, it might even be regarded as synonymous with ' educability,' the two terms being taken to imply the ability to acquire knowledge and to use it. Intelligence is thus the capacity to organize in a coherent manner the confused mass of experience which pours in upon us through the channels of the senses. Or we may even carry our analysis a step farther, and search for a general factor which enters into all the processes which are ordinarily described as ' intelligent.' To the writer it seems that attention constitutes such a factor, a belief which Binet himself appears to have held. The elementary intelligent processes—discrimination, comparison, analysis, synthesis—are all functions of attention. Intelligence would then be defined as " the functional efficiency of attention."

But, it may be urged, if there is such a general factor, why is not everybody equally good, or

equally bad, at everything? Why should not the great mathematician, for example, be an equally great historian, or zoologist, or literary critic? The writer believes the answer to be that those special aptitudes with which we are so familiar in ordinary life are in each case the consequence of a combination of general intelligence and special interest, the latter determining the particular channels into which the former is directed. For it is certainly a fact that good children, when young, are *generally* good all round, special tendencies becoming markedly apparent only during adolescence, when all kinds of interests and impulses hitherto latent begin to come into action. Moreover, a study of recent work in psycho-analysis will reveal to what a great extent the direction in which intelligent activity turns is determined by causes of which the individual concerned is quite unconscious.

There is, however, another experimental result at which the critics of mental tests frequently express incredulous surprise—namely, the fact that intelligence appears to cease developing

at something over seventeen years of age. But is this so incredible? Surely not when we remember the nature of intelligence. Intelligence is not knowledge, but the ability to organize and employ knowledge. Knowledge may grow indefinitely, but the growth of the ability to deal with knowledge may, like physical growth, cease comparatively early in life. The experienced, worldly-wise man of forty differs from the callow youth of sixteen or eighteen not in ability to handle experience, but in having a far greater amount of experience to handle, and thus to apply in meeting new situations.

One final word as to criticism: the critics should remember that the results reached by those engaged in work on mental tests are the product of careful and extensive experimental work carried out in accordance with the usual rules of scientific procedure. Hence *no* criticism can be effective unless it is based on experiments equally careful and extensive.

We have had occasion to draw a definite distinction between intelligence and knowledge

or attainment. The latter is a joint product of the former and environmental influences (of which, of course, teaching is one of the most important). But experience has shown that intelligence itself is native to the child—a gift, good or bad, at birth, which no means known to us can appreciably improve. The bright child remains bright, the dull child dull, to the end of the chapter. We cannot increase intelligence, but we can enable the child to make the *best* use of that amount of intelligence which he may happen to possess, and this is the true aim of all education.

Intelligence tests, then, are in no sense tests of teaching. They differ from the ordinary examination chiefly in three important ways: (1) they are directed to the estimation of the child's natural gifts, independently of such influences as home and social environment, and the effects of good or bad teaching; (2) they aim at enabling us to form a sufficiently accurate estimate of a child's educability before that child has, in fact, been educated to any great extent; (3) their results are expressed in

terms of a purely objective standard and are unaffected by any such subjective factor as the personal equation of a particular examiner.

With these aims in view the tests are so devised as to presuppose as little as possible beyond the fact that the child tested has been brought up (at whatever social level) under the ordinary conditions of a civilized community, and they are, of course, graded so as to be applicable to children of all ages and of all degrees of brightness. Their ability to predict (whereas the ordinary examination is in general really effective only after the event) should be, if proven, of the greatest value in enabling us to determine for each child the method of education fitted to make the best of his particular level of intelligence.

There are various methods of expressing degree of intelligence. One of the best known is by means of 'mental age' and 'intelligence quotient' (I.Q.). The mental age of a child is the age of the average child to whom he is equal in intelligence, and the I.Q. is the percentage ratio of the mental age to the actual

age. Thus a child of ten years may have a mental age of twelve years, his I.Q. then being 120. Experiment has shown that the I.Q. of a given child remains nearly constant (which makes it a very useful measure), so that the child in our example would, at the age of five years, have had a mental age of six years.

A second way of expressing level of intelligence is by means of ' percentile rank.' If a number of children be ordered according to degree of intelligence the percentile rank (P.R.) of any one of them is that percentage of the whole group which he just exceeds in intelligence. Thus if 70 per cent. of the children in the group are below the particular one selected in intelligence his P.R. will be 70. In view of what has been said of the nature of intelligence it is evident that the P.R. remains constant, for if a given child exceeds a certain percentage of all children in intelligence he will always exceed that percentage.

A third method, which is often very useful and significant, is to express degree of intelligence in terms of ' standard deviation.' If a group

of children of about the same age be tested, and the average or mean of their scores found, the deviation of each child's score from this mean is obtained and squared; the squares are then added and their mean found by dividing by the number of children. The square root of this mean is the standard deviation (S.D.). It is a measure of the 'scatter' of the intelligence of the group, the intelligence of a given child being expressed by saying that his deviation from the mean is so many times the standard deviation.

Other useful quantities in connexion with the scores of a group of children are the 'median' —that is, the middle score of the group (namely that made by the child whose P.R. is 50)—the 'lower quartile' (the score made by the child of P.R. 25), and the 'upper quartile' (the score made by the child of P.R. 75). Half the difference between the two quartiles is termed the 'semi-interquartile range' (S.I.R.), and provides another measure of scatter and one easier to obtain than the S.D.

Experiments in intelligence tests turn largely on the comparison of the order in which children

are ranged by such tests with the order in which they are ranged by other kinds of tests. The results of this comparison are expressed by means of what is termed a 'correlation coefficient.' It would be out of place here to enter into a description of the mathematical methods by which the degree of correlation is determined.[1] It may be stated, however, that any value of the correlation coefficient numerically greater than zero is presumptive evidence of the existence of a real connexion between the two sets of quantities compared; but for small values of the correlation coefficient the 'probable error' (P.E.) is so large as to make the result practically insignificant. As the magnitude of the correlation increases, however, the probability of the existence of a real connexion becomes rapidly greater, and when the value is about .5 the probability of *some* connexion begins to approach reasonable certainty. For values of the order .9 and upward the existence of a very close connexion is

[1] For such a description see, *e.g.*, Brown and Thomson, *The Essentials of Mental Measurement*, Part II.

practically certain. Perfect correlation is represented by the value 1.00.

In the descriptions of the experiments which follow the writer has been concerned primarily to illustrate the general principles involved in the examination of the validity of intelligence tests, and (if that validity is established) in the application of the tests for educational purposes. He has accordingly refrained from explaining in detail the more technical mathematical and statistical processes employed in obtaining some of the results.[1]

In actually carrying out the experiments the writer had two chief aims in view. In the first place, he wished to make the investigation not merely extensive, but also *intensive*. In other words, he was not concerned simply to get mass results, but also to make a close study of individual children. For it is clear that we can never be satisfied with the tests until we are confident that they are not only reliable for groups of children considered as wholes, but

[1] Such explanations may be found elsewhere ; *e.g.*, in Brown and Thomson, *op. cit.*, Part II.

are also reasonably certain to do justice to particular children. We require them to be at least as certain in this respect as the older methods of selection.

Secondly, the writer has relied as far as possible only on *objective* standards of comparison. In judging intelligence tests a method frequently employed is to compare their results with teachers' estimates. Now while no one would doubt the value of the estimate of a competent teacher, owing to the fallibility of human nature we can never feel quite the same certainty when a subjective factor, such as the personal equation of the teacher, enters into our calculations as we can when the standard of reference is purely objective. A man may be trained to estimate distances with some accuracy merely from observation, but we should never trust his estimate as we trust the foot-rule. In this connexion, therefore, it is necessary to state only that the writer, like others, has found close general agreement between teachers' estimates and results of the tests.

INTRODUCTORY

The chief problems attacked in the experiments to be described were the reliability of the individual test and of the group test as indexes of educability, and the methods of interpreting scores in group tests in such a way as to form an estimate of the mental capacities of the subjects tested—in particular, the methods to be adopted in the difficult case of adults and adolescents.

CHAPTER II

THE RELIABILITY OF THE STANFORD-BINET SCALE OF INTELLIGENCE TESTS AS AN INDEX OF EDUCABLE CAPACITY

INTRODUCTION. The problem of determining the use that can be made in educational work of intelligence tests in general, and of the Stanford-Binet scale in particular, reduces, as we have seen, essentially to this: Is it possible to foretell by means of these tests the limitations of a child's capacity for being educated, and hence to lay down the appropriate lines along which he should be taught?

The experiment herein described was an attempt to obtain decisive evidence of a quantitative nature on this point. Evidently the most direct method to adopt was to range a number of children in order of intelligence as determined by the S.B. scale, and then to arrange them in order of educable capacity by some entirely different (though sufficiently accurate) means

of estimation. The two orders could then be compared, and the correlation between them calculated.

For this purpose five batches of children (about twenty in each batch) were taken from five different schools. In order to ensure that the children had for some time had sufficiently favourable opportunities of learning, the schools selected were those known to be comparatively efficient, and the age range of the children was from ten to twelve, so that they had been at school for some years. In these circumstances the children's educable capacity could evidently be estimated by discovering what they had actually shown themselves capable of learning. As indexes of this capability arithmetic and English composition were chosen as being the most fundamental, as well as the most representative, subjects of the curriculum.

The children included members of both sexes, and they ranged in intelligence from exceptional superiority to definite mental defectiveness. In order to discount the effect of such heterogeneities as difference of sex, different methods

of instruction, etc., each batch was considered separately, in addition to the consideration of the group as a whole.

Every child was tested individually by the S.B. scale, and an order drawn up from the results.

In estimating the significance of what follows it should be borne in mind that extensive experiments in America—in the course of which children were retested under varying conditions, by different examiners, and at varying intervals—have shown that the intelligence quotient of a given child remains practically constant. Indeed, the average variation of I.Q., even over intervals of several years, is no more than about 5 per cent.[1] Hence, had the children in the experiment to be described been tested when entering, say, the senior school—*i.e.*, at about seven years of age—instead of at about eleven years of age, the order obtained would have been practically the same.

After being tested by the S.B. scale the

[1] *Cf., e.g.*, Terman, *The Intelligence of School Children* (Harrap), Chapter IX, especially pp. 138 ff.

children were set a paper in applied arithmetic. For this purpose the paper given on pp. 193 ff. of Dr Ballard's *Mental Tests* was used. On the child's performance in this test it was possible to assign to him an 'arithmetical' mental age, and hence to find his 'arithmetic quotient' (A.Q.) by a process similar to the calculation of the I.Q.

At this point it should be noted that the norms of performance taken differed somewhat from those given on p. 195 of *Mental Tests*. It became apparent early in the experiment that these were too low, and in fact Dr Ballard mentions the possibility of this on p. 195 of his book. The norms he gives are in arithmetical progression, and it therefore seemed probable that any increase due to a return of more favourable conditions of education would be marked by a multiplication of his norms by a constant factor (*i.e.*, the same for each norm at a given time). The first batch of results indicated an increase of about 25 per cent. in the norms, and this was strikingly borne out as further results came in.

In composition the children were required to write on *one* of the following subjects :

(1) How I spent my Easter Holiday.

(2) Describe what you have seen of the River X and its banks.

(3) Write a story called *Look before you Leap.*

(4) Write about this piece of poetry :

> Bright yellow, red, and orange,
> The leaves come down in hosts ;
> The trees are Indian princes,
> But soon they'll turn to ghosts.

As was expected, about 75 per cent. of the children chose the first subject, but this did not matter, as little difficulty was found in differentiating them.

In order to obtain an accurate standard a panel of five experienced examiners was formed, and every composition was marked independently by each examiner. A special method of marking was arranged. The examiners did not know the age of any of the children, and they were asked to mark a composition by assigning the child an age on it ; that is, a judgment of this kind was formed : " This composition is

about equal to that of the average child of x years of age." The mark x was then awarded. The mean of the five markings was found, and this was taken as the child's mental age as regards attainment in English composition. In awarding marks the examiners were asked to give primary regard to power of thought and to ability to express and arrange ideas in a logical and coherent manner. It may be remarked that, in the majority of cases, there was fairly close agreement between the examiners.

From the mental age thus obtained the 'composition quotient' (C.Q.) was calculated.

An order was drawn up by taking the mean of the two mental ages (for arithmetic and composition respectively) for each child, and another by taking the mean of the A.Q. and C.Q. (giving what may be called the 'educability quotient'—E.Q.). Evidently this last order is the significant one from the point of view of educability.

Results of the Experiment. In working out the correlations for the separate batches

Pearson's correction of Spearman's formula for ranks was employed (this was sufficiently accurate for correlations of the order obtained).

Calculated from the *quotients* the results were as follows :

	Group A (Boys)	Group B (Girls)	Group C (Boys)	Group D (Girls)	Group E (Boys)
Correlation between educability and general intelligence as measured by the S.B. scale	.93	.75	.88	.92	.91

Calculated from the *mental ages* the correlations were :

Group A	Group B	Group C	Group D	Group E
.93	.77	.89	.93	.92

It will be seen that these correlations are exceptionally high.[1] Moreover, the deviations from perfect correlations are almost entirely due to not more than about 15 per cent. of the children. For example, in Group B, where the

[1] It must, of course, be remembered that the high magnitude of the correlations here and in the sequel is in part due to the wide range of ability of the children tested, although this does not in any way vitiate the result. In the case of a group of children of approximately the same ability, such as a fairly uniform class, this homogeneity would alone tend to reduce the correlation, apart from any other factor.

correlations are lowest, the deviation was due very largely to three of the children, apart from whom the correlation coefficient would have been about .9. In these cases there were almost invariably obvious explanations of the discrepancy. Thus many of the cases were those of young, bright children who had not been given sufficient promotion, and therefore had had no opportunity of reaching a level of attainment appropriate to their mental age. Other causes of fall of E.Q. below I.Q. were late entry, bad attendance, inferior home conditions, serious emotional disturbance, physical defects, and malnutrition. Cases of marked rise of E.Q. above I.Q. are rare (amounting in this experiment to only about 4 per cent. of the whole) and seem to be generally attributable to a combination of mediocre or inferior intelligence with some temperamental characteristic such as capacity for hard work, or with home, school, or social conditions distinctly above the average. In such cases, however, it is probable that the superiority of E.Q. to I.Q. is not maintained to the same extent as the children

grow older. At all events, it is clear that the few discrepancies noted, far from detracting from the reliability of the S.B. scale, rather went to confirm its value when properly used.

At the other end of the scale converse results add further confirmation. For of the children of markedly inferior intelligence every one was placed in the attainment tests in practically exactly the position assigned to him by the S.B. scale. Hence the verdict of the latter—namely, that these children, even in favourable circumstances, were incapable of learning much—was completely vindicated.

One particular point is of interest in the last connexion. The norm of performance in the arithmetic test at eight years old is $\frac{5}{8}$ of a mark. Not a single child whose mental age, as given by the S.B. scale, was below eight years scored a mark in the arithmetic paper. The significance of this becomes clear in the light of the facts that (1) some of these children were actually twelve years old; and (2) the first problem in the arithmetic paper was as

follows : " If there are 100 apples on a tree and the wind blows down 17, how many are left on the tree ? " Some of the children diagnosed by the S.B. scale as mentally defective wrote pages of figures, but without scoring a single mark.

The group of 100 children was then considered as a whole. The rank correlation between the results of the educational tests and those of the S.B. test was obtained for the whole group. Calculated from the mental ages the correlation was .90, and from the quotients .89. The *absolute* (as distinguished from the ' rank ') correlation between the I.Q.s and E.Q.s of the members of the group was also calculated by the Bravais-Pearson product-moment formula, and was found to be .89. The probable error in these three correlations was of the order .01.

It will be seen from the table at the end of this chapter that the correspondence between I.Q. and E.Q. was, in the majority of cases, strikingly close. For the whole group the distributions of both I.Q. and E.Q.

approximated to normal. The constants of these distributions were :

I.Q.	E.Q.
Highest 149	Highest 138
Lowest 61	Lowest 64
Mean 104.36	Mean 102.33
Median 102	Median 102
Semi-interquartile range 13	Semi-interquartile range 9.5
Standard deviation 19.13	Standard deviation 14.56

It will be seen that the scatter of the E.Q.s is markedly less than that of the I.Q.s, probably indicating insufficient elasticity of promotion so far as the brighter children are concerned.

Lastly, it is important to remember that, for reasons already pointed out, the S.B. scale would have arranged these children in practically the same order had they been tested at (say) about seven years old instead of at about eleven. In other words, the order of intelligence given by the S.B. scale at about *seven years old* would correlate with the order of the children's educational attainment at *about eleven years old*

to nearly the same degree as the correlations given above. Hence, by testing at seven years old it would be possible accurately to predict educational attainments after some years' interval—that is, in effect, to form an exact estimate of educable capacity.

CONCLUSIONS. Here, then, we have a number of children of both sexes, educated in different schools, and of all levels of intelligence, and it turns out that their general intelligence, as assigned by the S.B. scale, correlates to a high order with their educable capacity as estimated from what they have actually shown themselves capable of learning in favourable circumstances. In about 85 per cent. of the cases the correlation is practically perfect, while in the remaining 15 per cent. there are generally clear reasons for the discrepancy.

It is true that, for many purposes, the number of children involved would not be sufficient to warrant the induction of a general principle; but it should be pointed out that the circumstances were such as to render the quantitative evidence afforded by this experiment practically

decisive in favour of the reliability of the S.B. scale as an index of educable capacity, especially when this evidence is taken in conjunction with the mass of qualitative evidence accumulated in America by comparison with teachers' estimates, following up of school careers, etc. For there does not appear to be any reason why this group of children should differ markedly from other groups of children, educated under sufficiently favourable conditions, in such a way as to modify the results of the experiment materially. In other words, there seems no reason to doubt that this group may be taken as a typical example of groups of children who have been afforded satisfactory opportunities of learning ; and, if this be granted, it evidently follows that conclusions true of this group would also be true, in their main principles, of any similar group.

There are also certain subsidiary conclusions of interest. Looking at the matter from the converse point of view, it follows that tests in arithmetic and English composition are good indexes of educable capacity, *always provided that*

(1) the tests are properly standardized; (2) the children to whom they are applied have been educated in favourable circumstances; and (3) due allowance is made for the ages of the individual children.

Attainment in arithmetic seems to correlate with general intelligence rather more closely than does attainment in composition, for whereas the respective correlations for the five groups of children between general intelligence and arithmetical attainment were .91, .78, .87, .92, .88 (calculated from the quotients), or .92, .83, .90, .93, .86 (calculated from the mental ages), the correlations between general intelligence and attainment in composition were .93, .42, .72, .84, .70 (calculated from the quotients), or .89, .32, .80, .82, .82 (calculated from the mental ages). Probably the lower correlation in the case of composition is partly real, partly apparent. Partly apparent, for, even with five independent markings, it is not possible to obtain an assessment of ability in composition of the same accuracy as the assessment of arithmetical ability. Partly real, for it seems

probable that attainment in composition depends on other factors (such as home and social environment) besides intelligence and education to a greater extent than does attainment in arithmetic. The reason for the comparative breakdown of the correlation as regards composition in Group B is not altogether clear.

If the S.B. scale is accepted as a reliable index of educable capacity children might well be tested by it twice during their school life, namely at about seven years of age, and again at about eleven years of age.[1] The first test would indicate the child's educability with sufficient accuracy to determine the methods according to which he should be taught in the senior school. The second test would serve to check the I.Q. obtained from the result of the first test, and would also settle the question of the child's capability of profiting by advanced instruction.

[1] For children of eleven, however, it will probably be necessary to use group intelligence tests, owing to the immense saving of time thereby effected.

Child's No.	I.Q.	E.Q.	Child's No.	I.Q.	E.Q.	Child's No.	I.Q.	E.Q.
1	67	71	37	91	104	72	96	98
2	141	138	38	112	113	73	101	100
3	101	107	39	98	107	74	96	100
4	126	124	40	95	91	75	122	112
5	73	69	41	85	88	76	110	105
6	132	125	42	111	114	78	95	97
7	61	64	43	142	130	79	102	107
8	112	100	44	99	89	80	101	102
9	109	105	45	92	91	81	102	97
10	119	98	46	116	108	82	89	88
11	125	108	47	83	81	83	89	92
12	81	100	48	114	106	85	131	118
13	123	113	49	90	92	86	115	110
14	78	81	50	91	84	87	71	82
15	121	112	51	80	83	88	138	130
16	75	67	52	88	87	89	128	120
17	122	120	53	96	89	90	121	112
18	102	97	54	92	96	91	128	120
19	117	111	55	133	107	92	129	119
20	139	125	56	112	106	93	102	107
21	108	114	57	120	110	94	102	115
22	107	100	59	111	121	95	105	104
23	71	75	60	130	108	96	107	108
24	113	108	61	92	105	97	107	100
27	92	101	62	86	94	98	115	110
28	144	125	63	110	102	99	108	103
29	149	117	64	113	117	100	88	95
30	108	99	65	102	107	101	96	82
31	103	101	66	67	76	102	78	98
32	92	92	67	89	95	103	91	93
33	93	90	68	128	121	104	99	98
34	113	102	69	84	94	105	94	94
35	82	102	70	78	92			
36	117	114	71	134	134			

The missing numbers are those of children who were absent from some of the tests.

CHAPTER III

THE DERIVATION OF MENTAL AGES FROM SCORES IN A GROUP TEST

IF a number of children have been examined by means of a group intelligence test we are able, by comparing their scores, to arrange them in order of relative intelligence. This is sufficient for the important purpose of classifying the children according to the degree of development which their intelligence has reached. But it is quite inadequate for other, and equally important, purposes. For we cannot compare the real degrees of intelligence of the children until we have taken their ages into account. For example, one child may score more than another in the test, but the first child may be fourteen years of age, while the second may be only ten. Thus, although the intelligence of the first has developed more than that of the second (so that he would be classified higher), the latter may in *relative* development be ahead of the former. That is, the ten-year-old may be

more developed than was the fourteen-year-old at ten years of age. Therefore, before we can be in a position to judge as to which child is really the more intelligent we must somehow manage to reduce them to a common denominator by taking their ages into account.

The process we have indicated is carried out (as we have seen in another connexion) by assigning to each child a mental age, dividing this by the actual age, and multiplying by 100 to obtain the I.Q. The question then arises as to how mental age is to be obtained from score in the test. The mental age is the age of the ' average ' child to whom the given child is equal in intelligence. It will therefore be necessary to determine the score of average children in the test, or, to put it another way, to calculate the average scores in the test of children in various age groups. Thus we may take all the ten-year-olds tested, and find their average score ; similarly with the eleven-year-olds, and so forth. But if our results are to be valid the age groups must be sufficiently large and sufficiently random to constitute fair samples.

Having obtained these average scores or 'age norms' we proceed to find the mental age of any given child as in the following example: Suppose a child of ten scores 60, and we find that 60 is the average score of the twelve-year-olds. Then the child's mental age, *as measured by the test*, is twelve, and his I.Q. is 120.

But is the mental age, *as measured by the test*, the *true* mental age, *i.e.*, does it express with sufficient accuracy the child's standing relatively to the average child? For the aspects of intelligence are so many and various that it is not easy to devise a test to measure them all, and it is possible that we may miss some point in which the child shows up particularly well.

Now unfortunately it is possible to assign a mental age only by means of some test, and no test is perfectly comprehensive. What is needed, then, is a specially comprehensive test which, though not perfect, will serve as a standard by which the results of other tests may be judged. Probably the most accurate and comprehensive scale of tests at present available, at any rate

for children up to about fourteen years of age, is the Stanford-Binet scale. It will therefore be important, after deriving mental ages from the group test in the way indicated above, to check these mental ages by comparison with the corresponding Stanford-Binet mental ages in the case of children who have been tested both by the group test and by the Stanford-Binet scale.

We may now proceed to illustrate the principles we have laid down by an example from actual practice.

472 children, ranging from ten to thirteen years of age, were tested by Form A of the *Terman Group Test of Mental Ability*, for which no age norms of performance had at that time been published. After the examination the booklets were scored and the average scores or age norms for different age groups calculated. A 'curve of performance' was then obtained by plotting average scores for various ages against those ages.[1] Curves of performance in

[1] This procedure will give valid results only if the distribution of scores in each age group is approximately normal—as it actually was in the experiment here described.

group tests of this kind usually approximate, over a considerable range, to straight lines,[1] and the curve obtained in this case was no exception. The age norms found were :

Age in years	$10\frac{1}{4}$	$10\frac{3}{4}$	$11\frac{1}{4}$	$11\frac{3}{4}$	$12\frac{1}{4}$	$12\frac{3}{4}$	$13\frac{1}{4}$
Average score	48	56	65	71	82	83	88

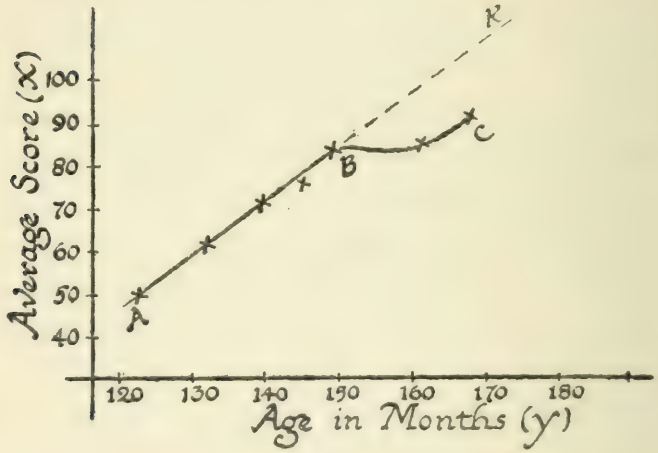

FIG. 1

Hence the curve of performance was as in Fig. 1.

[1] The true curve of performance is probably an ogive (*cf.* footnote on p. 65). The straight lines obtained as described in this chapter are simply intended to be close approximations to a large portion of the ogive. The simple nature of the equation to the straight line renders it very convenient to use in practice.

The section AB is a straight line, the points corresponding to ages 10¼, 10¾, 11¼, 12¼ years respectively lying practically exactly on this line, while the point corresponding to 11¾ years is very nearly on it. As to the eccentric section BC a word of explanation is necessary. After the age of about 12 years many of the brighter children have left the elementary schools for secondary and other schools. Hence the age groups above 12 become less and less representative samples, and their performances will drop below what might have been expected from a consideration of the performances of the younger children. We must therefore neglect the portion BC of the curve of performance, and replace it by the dotted line BK, which is an extension of the straight line AB. The higher mental ages should then be obtained, not from BC, but from BK, which represents the average scores of higher age groups on the assumption that intelligence continues to develop at the same rate as for the younger age groups.[1]

If we call the mental age y and the score in

[1] This point will be elucidated more fully in Chapter IV.

the group test x, the straight line ABK is represented by an equation of the form

$$y = mx + c$$

which is the typical equation for a straight line.

For ABK it can easily be shown that, if y be expressed in *months*, $m = \frac{7}{10}$ and $c = 90$. Hence

$$y = \tfrac{7}{10}\, x + 90$$

that is, the mental age can be derived from the score by the following formula :

Mental age in months = ($\frac{7}{10}$ of score) + 90

Of course, once having obtained the straight line ABK, it would be possible to read off from it directly the mental age corresponding to any given score without working out the formula. Also a table might be drawn up from the line or from the formula giving the mental age for various scores. Had the curve of performance not been a straight line its equation would have been more complicated, and the determination of the formula might have presented some difficulty. In such a case it would be more convenient to construct a table direct from the curve.

So far so good—but a fresh point now arose.

It was known that the batch of children tested was likely to be, on the whole, above the average of the total population. Hence this group of children set a standard above the average, so that the mental ages given by the above formula would be lower than the true mental ages (*i.e.*, the mental ages obtained relatively to a fair sample of the *whole* population, for such a sample would set a lower standard than the particular group of children tested). Therefore the formula could be regarded only as a first approximation to the truth, good enough for arranging these particular children in their order of intelligence, but needing to be checked and modified where necessary if it was to be used for obtaining true mental age.

The formula was checked in the following way : One hundred children were tested individually by the Stanford-Binet scale, and were then given the Terman group test. A curve was plotted giving the relation between score in the group test and S.B. mental age. As was expected, this curve gave somewhat higher mental ages for given scores than did the line

ABK in Fig. 1. The number of *high* scores was too few to form a basis for safe generalization, but in the higher ranges there was an apparent tendency for the new curve and the line ABK to converge toward one another. This looked as if the formula obtained from ABK was more accurate for high scores than for low.

Another check became available later. In a more recent edition of the *Manual of Directions* for the Terman group test[1] there appeared a table, based on 306 cases, giving tentatively the probable correspondence between scores and S.B. mental ages. The line plotted from this table was found to agree very closely indeed in the middle ranges with the curve obtained from the 100 cases mentioned above. But for the lower ranges the curve gave rather lower mental ages than the new line. It is not improbable, however, that the curve is more trustworthy here, for Terman mentions in a footnote that the mental ages given in his table are likely to be somewhat too high.

[1] *Cf.* p. 10 of the *Manual of Directions* in the 1921 American edition, not yet obtainable in England.

The relation between the original line ABK, the curve obtained afterward by comparison with the S.B. scale, and, finally, the line plotted

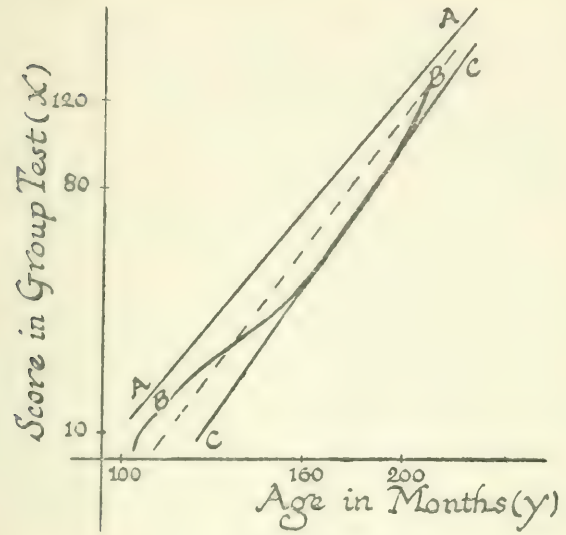

FIG. 2

from the data given by Terman is represented in Fig. 2, which is drawn only approximately to scale.

In Fig. 2 AA represents the performance of the 472 children in the first experiment (ABK in Fig. 1), BB the curve afterward obtained by

the writer from the comparison of the scores of the 100 children with their S.B. mental ages, and CC the line plotted from Terman's data referred to above.

The following points will be noted:

(1) The convergence of AA, BB, and CC as the higher scores are approached.

(2) The coincidence of BB and CC in the middle range.

(3) AA and BB both give lower mental ages for the lower scores than CC.

Considering AA, BB, and CC in conjunction, and allowing so far as possible the appropriate weight to each, we are led to a line in the position of the dotted line in the figure as probably a close approximation to the true representation of the relation between mental age and score in the Terman group test.

When this line was plotted and its equation expressed in the form

$$y = mx + c$$

it was found that, compared with the original line AA, m had diminished from $\frac{7}{10}$ to about $\frac{6}{10}$ or $\frac{3}{5}$, while c had increased from 90 to about 110.

Thus the new equation will be very nearly

$$y = \tfrac{8}{5}x + 110$$

so that mental age may be obtained from score by the formula

Mental age in months = ($\frac{8}{5}$ of score in Terman group test) + 110

Notice that for small scores the first item on the right will be unimportant compared with the second term—the constant 110; while as the score increases the first term becomes more and more important. Thus the lower mental ages assigned will depend very largely on the value given to c (which we have taken as 110), while the higher mental ages will depend more or less equally on c and m (which we have taken as $\frac{8}{5}$).

Now the value of m depends on the slope of the line, and it will be seen from Fig. 2 that AA and CC do not differ greatly in slope, while the slope of BB, though it varies somewhat from point to point, is on the whole much the same as that of AA and CC. For AA $m = \frac{7}{10}$, for CC m = about $\frac{5 \cdot 6}{10}$. Therefore there is unlikely to

be a serious error in the value we have taken for m, namely $\frac{8}{5}$. The value of c, on the other hand, depends on the value of y for which x vanishes; that is, it represents the mental age in months for which the score becomes zero—the age below which the average child cannot even begin to do the test. For AA $c = 90$, a mental age of $7\frac{1}{2}$ years. For CC c = about 116, a mental age of 9 years 8 months. There is a considerable discrepancy between these two values of c, and we should therefore find somewhat serious differences in the lower mental ages according as we adopted one or the other.

But it is possible to get nearer to the true value of c. For it was pointed out that the formula $y = \frac{7}{10}x + 90$ gives mental ages which are too low, and we saw that the error increases as the score diminishes. On the other hand, while the mental ages found from CC are likely to be too high, the comparison here is at any rate one between score and true or S.B. mental age. The value 116 for c is thus indicated as being considerably nearer the truth than the value 90. Moreover, the curve BB points to a value for

c of about 106, or a mental age of 8 years 10 months. It is therefore improbable that there is a serious error in the value we have provisionally decided to take for c, namely 110, or a mental age of 9 years 2 months. But it will be clear that the precise determination of this constant is important.

While the formula $y = \frac{3}{5}\ x + 110$ probably gives with some accuracy the correspondence *in general* between true mental age and score in the Terman group test, it must not be forgotten that errors may occur in individual cases. If the tests were perfect, children testing at the same mental age on the S.B. scale should score the same in the group test. But in practice it not infrequently happens that children of the same S.B. mental age make considerably different scores in the group test. The reason is not far to seek. In the group of 100 children who were given both tests all cases where equal or nearly equal mental ages were correlated with different group test scores were carefully analysed. In nearly all such cases it at once became evident that the discrepancies were due to the fact

that the two tests did not cover exactly the same ground. Certain aspects of intelligence to which some of the S.B. tests are directed are not touched by the Terman group scale. For example, that function of intelligence which determines control of visual imagery is not specifically tested by the latter, while in the S.B. scale there are at least two tests (No. 6 in year xiv, and No. 2 in ' superior adult ' group) which depend primarily on the manipulation of visual imagery. Thus, if a child made a lower score than might have been expected from his S.B. mental age, it was found on analysis that he had shown up well when tested by the S.B. scale in some direction, such as control of visual imagery, which is not specifically tested by the Terman group test.

On the whole, however, the distribution of points obtained by plotting the S.B. mental ages against the group test scores shows that the S.B. mental ages of two children with the same score are much more likely to agree fairly closely than to differ widely. But, at the same time, it is clear that group tests should be

devised in such a way as to make them as comprehensive as possible. Otherwise in individual cases different scales may give considerably different results; and we cannot feel complete confidence in the scientific validity of our scales as trustworthy instruments of measurement with reference to an objective standard until we have advanced to a point where different scales, when applied to the same child, will give results in close agreement.[1]

[1] The writer may perhaps be permitted to refer here to the *Simplex Group Intelligence Scale* (Harrap), which he has recently published, and which is devised to form a scale sufficiently comprehensive as regards both the various aspects of intelligence which it covers and the age range of the children to whom it may be suitably applied. The Simplex scale was administered to about a hundred children whose mental ages had been previously estimated by several other methods (including the S.B. individual test). The correlation between these mental ages and the scores in the Simplex test was found to be .94. There were only three marked discrepancies, a proportion so small that it might well have been due to accidental causes. Neglecting these three cases, the correlation rose to .97, which is very high indeed.

CHAPTER IV

METHODS OF ESTIMATING THE TRUE INTELLIGENCE QUOTIENTS OF ADULTS AND ADOLESCENTS

THE results of experiment seem to show that intelligence grows at an approximately constant rate up to the age of something over fourteen years. We may term this age the ' critical age.' Beyond the critical age the growth of intelligence slows down until it comes to a complete stop at the age of something over seventeen years. The age at which this cessation of growth occurs will be called the ' terminal age.'

As we have seen, a convenient way of expressing the absolute level of intelligence of a given child at a given time is by means of ' mental age.' The rate of growth of intelligence will then be the ratio of the mental age to the actual age. This ratio, expressed as a percentage, is the ' intelligence quotient,' and is found to be

approximately constant for the same child from early childhood up to the critical age, though different for different children. The I.Q. (which really represents the rate of mental

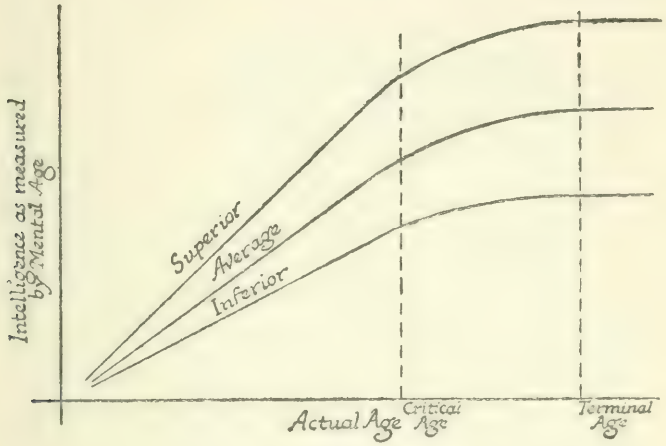

FIG. 3

growth) is thus a highly convenient index of a child's degree of intelligence. Fig. 3 represents the growth of intelligence of three children, one inferior, one average, and one superior.

It is by no means certain, though for convenience it has been assumed in Fig. 3, that the

critical and the terminal ages are approximately the same for all children.[1]

The child's degree of intelligence is, then, given by the I.Q., which represents the rate of growth during the period when that rate remains nearly constant, namely below the critical age. But a difficulty arises when a subject is not tested until the adolescent or adult period, for, owing to the slowing down of mental growth, the ratio of mental age to actual age will change, and will therefore no longer be an expression of the true I.Q., but will enable us only to set certain limits to the subject's intelligence.

The problem before us, therefore, is as follows : How, from the score obtained in a mental test by an adult or an adolescent subject, can that subject's *true* I.Q. be estimated—*i.e.*, the I.Q. which would have been found for him had he been tested in childhood ?

It will be simpler to develop the methods of

[1] Even should it be found ultimately that in the case of all children intelligence ceases to grow at approximately the same age, different children will have reached at that age very different levels of mental development. *E.g.*, the dull or defective child may have attained a mental age of only ten or twelve years, at which he remains for the rest of his life.

solution proposed by reference to a particular group test for which the age norms of performance are known. For this purpose we shall take the Otis group test. All that will be said would

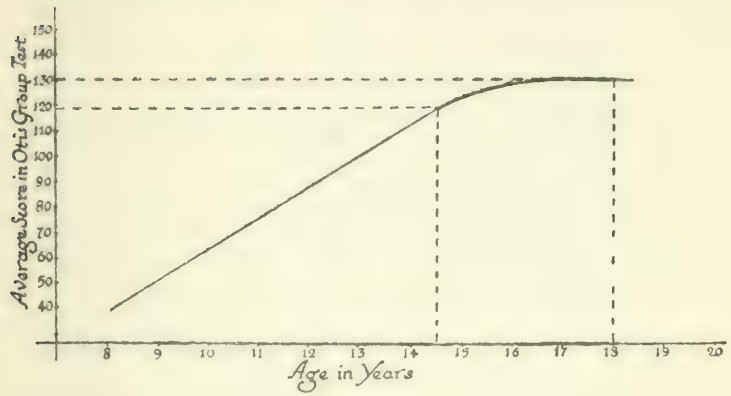

FIG. 4

be equally applicable, *mutatis mutandis*, to any other point-scale test.

I. A DIRECT METHOD. Fig. 4 represents the curve of performance in the Otis group test, as plotted from the table of norms.[1] This curve is a rough reflection of the curve of growth of intelligence of the 'average' child. A given child's mental age is found by taking his score in

[1] See *Manual of Directions* for Otis group test, p. 65 (Harrap).

the test and finding from the curve the age to which it corresponds. The percentage ratio of this age to the child's actual age will give his I.Q. A glance at the curve in Fig. 4 will show that the critical age is between 14 and 15 years (actually 14 years 5 months—see table of norms) and the terminal age about 18 years. Below the critical age the improvement in score of the *average* child is 1 point per month, or 12 points per year.

Let us first consider the case of a fairly bright child below the critical age.[1] Suppose at 12 years of age he scores 102. His mental age is then 13 years 2 months, and his I.Q. 110. From 12 to 14 years his intelligence will continue to grow at much the same rate, as he is below the critical age. At 14 years his mental age should therefore be 15 years 5 months. His rate of improvement of score will be 10 per cent. above that of the average child (*i.e.*, it will be 1.1 points per month), and if he is again tested at 14 years his score will be $102 + (24 \times 1.1)$—that is,

[1] In what follows, the not improbable assumption is made that the intelligence of a given child continues to bear to the intelligence of the average child nearly the same relation after the critical age as before that age, though the rate of growth of intelligence of both children will then be diminishing.

about 129. From the curve in Fig. 4 the age corresponding to a score of 129 is about 17 years, and the ratio of this to 14 years will thus be greater than the true I.Q. The reason for this discrepancy is, of course, that during the period 12 to 14 years the child's mental growth has not been following a curve which has been rising less and less rapidly like the one in Fig. 4 for a part of the time, but has still been increasing at a comparatively steady rate. Thus, to obtain the true I.Q. we must extend the straight portion of the curve in Fig. 4. The age along this straight extension corresponding to a score of 129 will be 15 years 5 months, and the ratio of this to 14 years gives 110, the true I.Q. In other words, when the child's mental age rises above the critical age, we must compare him with a fictitious average individual[1] whose intelligence (and therefore mental age) continues to grow at a steady rate indefinitely, if we are to measure the true I.Q. The mental age thus obtained we shall call the ' effective mental age.' When the actual mental age is below the critical age it

[1] *Cf.* also *Manual of Directions* for Otis group test, pp. 53 ff.

will evidently be identical with the effective mental age. Clearly this method of finding the effective mental age is applicable to a subject of any age.

The above case points the way to the solution of the main problem we are considering—the determination of the true I.Q. of the adolescent or adult. In such cases a difficulty arises additional to that found in the case of the bright child just below the critical age. For what are we to take as the denominator in the I.Q. ratio for subjects over the critical age and eventually over the terminal age?

We may consider the case of the terminal age first. Let us now suppose that intelligence, instead of slowing down after the critical age and gradually coming to a final stop at the terminal age, continues to grow at the same constant rate as below the critical age, and then comes to a *sudden* stop at the same level as is reached in actual life at the terminal age. Thus we must take the score of the average subject in the test as continuing to increase by 1 point per month. Now from the curve in Fig. 4 we see that, for the

average individual, the ' terminal score '—*i.e.*, the score reached at the terminal age—is 130. Had the score still increased after the critical age of 14 years 5 months (at which the score is 117) at the rate of 1 point per month, the score of 130 would have been reached at the age of 15 years 6 months. We may call this age the ' effective terminal age.' Then for all subjects older than the terminal age we must take the effective terminal age as the denominator of the I.Q. ratio. The I.Q. for such subjects will therefore be the ratio of the effective mental age (calculated as previously described) to the effective terminal age.

It still remains to deal with the case of those subjects who are younger than the terminal age, but older than the critical age. The procedure here is to find the age at which the score corresponding on the curve in Fig. 4 to the actual age of the subject concerned would have been reached if the score of the average individual continued to increase by 1 point per month after the critical age. The age thus obtained may be termed the ' effective age,' and must be used as the denominator of the I.Q. ratio for the subject

concerned. For subjects below the critical age the effective age will evidently be identical with the actual age, while for subjects above the

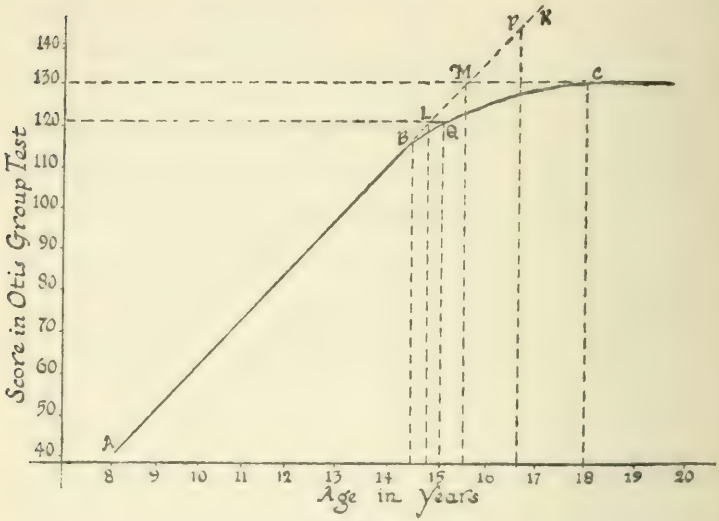

FIG. 5

terminal age it must be taken as equal to the effective terminal age.

The foregoing may perhaps be made clearer by an example illustrated by the diagram given in Fig. 5.

The curve ABC in Fig. 5 is the curve of performance in the Otis test as in Fig. 4. BK

is the extension of the straight portion AB of the curve. B is at the critical age, C at the terminal age, and M at the effective terminal age.

Suppose a child of just 15 years makes a score of 145. The point P on BK gives the effective mental age as 16 years 9 months. The point Q on the curve ABC corresponds to the actual age of 15 years, at which the average score is 121. QL is parallel to the age axis, and the point L on BK gives the effective age as 14 years 9 months. Thus the true I.Q.—that is, the percentage ratio of the effective mental age to the effective age—is $\frac{16 \text{ years } 9 \text{ months}}{14 \text{ years } 9 \text{ months}} \times 100$, *i.e.*, $\frac{201}{177} \times 100$, or 114.

The procedure for finding the mental age of an adult or adolescent is therefore briefly as follows:

From the subject's score obtain the effective mental age from the straight line ABK. From the curve ABC obtain the score corresponding to the subject's actual age, and hence from the straight line ABK the effective age corresponding to this score.[1] The table of norms can, of

[1] Both theory and practice show that the most probable form of

course, be used instead of the curve, provided we remember that in finding the effective age and the effective mental age the average score must be supposed to increase at the same rate above the critical age as below that age. The true I.Q. will then be given by the formula $\frac{\text{effective mental age}}{\text{effective age}} \times 100$, which is more general than the older formula $\frac{\text{mental age}}{\text{actual age}} \times 100$, for the latter is applicable only when the mental age and the actual age are both below the critical age.

The foregoing cannot be applied, except in a modified form, to such scales as the Stanford Revision of the Binet tests. In the latter scale Terman takes the effective terminal age as 16 years, and arbitrarily fixes the value in months of the tests in the 'average adult' and 'superior

the curve of performance in a group test is the curve known as an 'ogive.' If, however, an ogive is much drawn out it approximates closely to a straight line over a considerable range. This is what usually happens, as in the Otis test. If the data permit the plotting of the ogive, intercepts for finding effective mental age would then be made on the latter instead of on the simple extension of the nearly straight portion; but the difference will be practically important only for scores approaching the maximum possible. It should be noticed that the drop in the ordinary curve of performance after the critical age is due to the slackening growth of intelligence, and not to the tendency toward the ogive form.

adult' groups in such a way as to cause adults to test *on the average* at a mental age of 16 years. The method outlined here could thus be applied to the Stanford Revision only after an investigation had been made to determine the actual average performance of subjects from about 14 to (say) 20 years of age in the upper groups of the scale, and a curve of performance plotted from the results of this investigation apart from any arbitrarily predetermined value in months for the harder tests.

To test satisfactorily the method that we have been discussing it would be necessary to retest the same children at intervals during adolescence in order to discover whether the I.Q. calculated by this method remained nearly constant for the same child and equal to its value below the critical age. Unfortunately, such retests with the Otis scale are not at present available, but, in their absence, artificial cases may be invented by the following device :

A child of given age and given score at that age is assumed. Knowing this, the distribution tables at the end of the Otis *Manual of Directions*

will enable the child's percentile rank (P.R.) to be approximately calculated. On the basis of this P.R., examination of the distribution in the other age groups will make it possible to estimate this same child's probable score at other ages, since his P.R. remains constant, due allowance being made for the fact that the higher age groups are not fully representative. Thus, for example, if his P.R. is 70, to find his probable score at 12 years of age look at the distribution of scores for the age group 12, and find the score of P.R. 70.

Two examples may be given in illustration, one of a bright and one of a dull child.[1]

(1) A bright child who at 16 makes a score of 180. The probable scores of this child for various ages were found, and hence his effective mental ages, which, with the effective ages, gave the values of his I.Q.

The results are given in the table on p. 69.

Hence for this case the method gives a nearly constant I.Q. Note that by the ordinary

[1] These two examples were suggested to the writer (by Mr Burt) as a test of his method.

method this child would at 12 be supposed to have a mental age of 18 (130 being the average score at 18), giving an I.Q. of 150, very wide of the mark. Above 12 years it would not

Actual Age of Child (Years)	Probable Score	Effective Mental Age (Months)	Average Score at Child's Actual Age	Effective Age (Months)	True I.Q.
18	182	238	130	186	$\frac{238}{186} \times 100 = 128$
16	180	236	125	181	$\frac{236}{181} \times 100 = 130$
15	173	229	122	178	$\frac{229}{178} \times 100 = 129$
14	160	216	112	168	$\frac{216}{168} \times 100 = 129$
12	130	180	88	144	$\frac{180}{144} \times 100 = 132$
10	104	160	64	120	$\frac{160}{120} \times 100 = 133$

be possible to assign him a mental age by the ordinary method.

(2) A dull child of I.Q. about 70. This child's mental age when 12 years old would then be about 8 years 5 months, and therefore his score would be 45 (see table of norms).

The results are given in the table on p. 70.

Here the variation of I.Q. is too great to be satisfactory, but it should be noted that the

probable scores obtained for this child by the P.R. method for ages above 14 are certainly considerably higher than his scores would actually be in practice. For the probable scores

Actual Age of Child (Years)	Probable Score	Effective Mental Age (Months)	*Average* Score at Child's Actual Age	Effective Age (Months)	True I Q.
10	30	86	64	120	$\frac{86}{120} \times 100 = 72$
12	45	101	88	144	$\frac{101}{144} \times 100 = 70$
14	68	124	112	168	$\frac{124}{168} \times 100 = 74$
15	81	137	122	178	$\frac{137}{178} \times 100 = 77$
16	90	146	125	181	$\frac{146}{181} \times 100 = 80$
18	99	155	130	186	$\frac{155}{186} \times 100 = 83$

make his rate of increment of score after 14 *greater* than that of the average child, which would clearly not be the case. Hence in practice our method would give an I.Q. constant within limits considerably narrower than is apparent from the example. It simply happens that the P.R. method, which is here an artificial device dependent for its accuracy on the accuracy and comprehensiveness of the distribution

tables, does not, in this case, give probable scores sufficiently near to the true scores.

Now let p stand for the child's score and q for the *average* score at his age. Then it will be seen from the above examples that his effective age is $q + 56$, and his effective mental age $p + 56$, so that his true I.Q. is $\frac{p+56}{q+56} \times 100$. Similarly, for the Terman group test we should have, in consequence of the formula obtained in the last chapter, true I.Q. $= \frac{\frac{3}{5}p + 110}{\frac{3}{5}q + 110} \times 100$. Evidently it would be quicker and simpler to find the true I.Q.s from formulæ like these than direct from the curve as in Fig. 5. But the introduction and discussion of the latter was necessary to make the principle of the method clear. As a matter of fact, in practice, once the norms have been established (and hence the critical and terminal ages) for any group test, the best thing to do would be to construct, by means of a formula like the above, a ready-reckoner giving true I.Q. direct from score and actual age.

The results we have arrived at may perhaps

be best summarized by a recapitulation of the definitions of the new terms used.

The *critical age* is the age at which intelligence ceases to grow at an approximately constant rate, and its development begins to slow down.

The *terminal age* is the age at which intelligence stops growing altogether.

The *terminal score* is the average score reached at the terminal age.

The *effective terminal age* is the age at which the terminal score would be reached if intelligence continued to grow after the critical age at the same rate as before the critical age.

The *effective mental age* of a subject is the age at which the score he makes would be reached by him if his intelligence continued to grow indefinitely at the same rate as below the critical age. When the effective mental age is less than the critical age it is identical with the actual mental age.

The *effective age* of a subject is the age at which the *average score* corresponding to his actual age would be reached if intelligence continued to grow at the same rate after the critical age as

before that age. If the effective age is less than the critical age it will be identical with the actual age ; if it is greater than the terminal age it must be taken as equal to the effective terminal age.

The general formula for the I.Q. then becomes the percentage ratio of the effective mental age to the effective age. This formula gives the *true* I.Q. for adults and adolescents—that is, the I.Q. which characterized them when they were below the critical age.

For the Otis group test the critical age is 14 years 5 months, the terminal age 18 years, and the effective terminal age 15 years 6 months. It will be important to determine whether the critical, terminal, and effective terminal ages are (*a*) approximately the same for different children, (*b*) approximately the same for different test scales.

II. An Indirect Method. The writer has also obtained estimates of true I.Q.s by means of a percentile rank method. This method rests on two facts :

(1) The *percentage* distribution of I.Q.s in any particular age group, if sufficiently large, will

approximate closely to the percentage distribution of I.Q.s for the whole population. This simply means that any two large age groups will probably have about the same distribution of I.Q.s.

(2) Within a given age group the distribution of I.Q.s will be the same as that of the mental ages, since the denominator of the I.Q. ratio will be approximately the same for every individual in the group. The narrower the limits of age between which the group is taken, the nearer will the agreement be.

In order to use this method two things must be known :

(*a*) The general distribution of I.Q.s below the critical age (these being true I.Q.s).

(*b*) The distribution of scores (and therefore of mental ages) in an age group with a narrow age range about the age of the particular individual who is being dealt with.

The first is now fairly well known, as is also the second in the case of some group tests.

The method, which is simple enough to apply, needs only a brief description. Knowing the

actual age of the subject concerned, and the score he has made, find from the distribution tables the P.R. of his score, and therefore of his mental age, in an age group of individuals of nearly the same age as himself. As pointed out above in (2), the P.R. of his mental age within this age group will be the same as the P.R. of his true I.Q. in the age group, and the last is the same as the P.R. of his true I.Q. when he was in any other age group (see (1) above). Now the distribution of true I.Q.s is known from the age groups below the critical age. Hence, knowing the P.R. of the given subject, his true I.Q. can be found at once.

For example, suppose a person of something over 16 years of age makes a score of 162 in the Otis test. His P.R. in the 16–17 age group, found from the distribution table at the end of the Otis Manual, is about 75. Now in the general distribution of true I.Q.s the 75 percentile is about 109. Hence this person's true I.Q. is about 109. For greater accuracy, however, the age ranges should be narrower than those in the Otis distribution tables.

The method we have been discussing is applicable, with slight modifications, to the Stanford-Binet individual scale, as explained in the appended note.

NOTE ON A METHOD OF ESTIMATING THE TRUE STANFORD-BINET I.Q.s OF ADULTS

The following method of estimating the true Binet (Stanford Revision) I.Q.s of adults, or of children approaching 16 years of age, which has occurred to the writer in the course of his general inquiry into methods of estimating the true I.Q.s of adults and adolescents, may perhaps be of interest to those engaged on Binet tests.

The difficulty which occurs in this connexion with the Binet scale lies in the fact that, owing to lack of extensiveness in the upper ranges, I.Q.s of adults, as measured by the scale, are bound to drop below their true values. The greatest possible I.Q. at which an adult can test is about 126.

For the method to be used two things must be known: (1) the distribution of children's

Binet I.Q.s ; (2) the distribution of the mental ages of adults as measured in the ordinary way by the Binet scale.

The first is fairly well known (*cf.* Terman, *The Measurement of Intelligence* (Harrap), pp. 66 and 78), but there are not yet sufficient data to obtain the second accurately, though Terman gives a diagram based on 62 ' normal ' cases on p. 55 of the book referred to. The writer has, however, found the latter useful as a rough working basis. (*N.B.* In Terman's diagram on p. 55, 15.11 seems to be a misprint for 16.11.)

To find true Binet I.Q.s proceed thus : Having obtained the mental age of any particular adult in the ordinary way, find his P.R. from the distribution table for mental ages of adults. Since, in finding the I.Q., the actual age is taken for every adult as 16 years, the P.R. obtained will also be the given person's P.R. as regards I.Q. Then from the distribution tables for children's I.Q.s, which are true I.Q.s (the distribution remaining the same for different age groups), it is possible to find directly the approximate value of the given

adult's true I.Q., his P.R. having been found as explained above.

In conclusion, two points should be noted: (1) for I.Q.s below about 105 the value obtained in the ordinary way is probably very nearly the true one; and (2) Terman's distribution diagram for adults can be used only up to mental ages of about 18 years 6 months. When more extensive data are available for obtaining this distribution the percentile rank method should give I.Q. values of considerable accuracy.

TABLE OF PROVISIONAL VALUES OBTAINED BY P.R. METHOD

Mental Age found by Usual Method	True I.Q.
16 years 6 months	107
16 ,, 9 ,,	108
17 ,, 0 ,,	110
17 ,, 3 ,,	112
17 ,, 6 ,,	115
17 ,, 9 ,,	118
18 ,, 0 ,,	122
18 ,, 3 ,,	126
18 ,, 6 ,,	132

CHAPTER V

THE RELIABILITY OF THE GROUP INTELLIGENCE TEST AS AN INDEX OF EDUCABILITY

IN Chapter II we discussed an experiment the results of which pointed to the fact that the Stanford-Binet scale provided an accurate means of gauging the educable capacity of a child. This scale is, however, administered individually, and thus involves the expenditure of a considerable amount of time, especially when large groups of children are to be tested. In practice, therefore, it is necessary to employ group tests which can be applied to a great many children simultaneously, retaining the individual scale as a standard of reference and a source of further information in particular cases.

The question therefore arises as to the reliability of the group test as an index of educability. There are two methods of approaching the problem, one indirect, the other direct. The first method consists in comparing the results of the group test with those of the

individual test, the reliability of which has already been largely established. The second method consists in comparing the results of the group test with those of ordinary educational tests which may be taken as affording reasonably accurate estimates of the educable capacities of children who have been at school for some years. Experiments illustrating both methods will now be described.

I. The Indirect Method. The group of children tested individually by the S.B. scale, as described in Chapter II, was afterward given a typical group intelligence test. It should be noted that the group tests at present in vogue, though they differ from one another in detail as regards both particular forms of tests employed and particular *items* of different tests, conform largely to a type. If the reliability of such a typical group test can be established, it therefore follows that any group test of similar type, provided it is reasonably comprehensive, and provided it has been shown to be sufficiently suitable by trial, is likely to be equally reliable.

In the case with which we are dealing the figures obtained from the comparison of the results of the individual and the group tests were considered for each batch of children separately (see Chapter II) as well as for the group as a whole.

The rank correlations for the five batches between scores in the group test and mental ages obtained by the individual test were as follows :

Group A	Group B	Group C	Group D	Group E
.86	.86	.91	.88	.92

When the group was taken as a whole the rank correlation worked out at .89.

It will be seen that these correlations are high, and it may therefore be safely concluded that the group test is, on the whole, a sufficiently reliable index of educability.

On examining the results in detail two conclusions were drawn :

(1) The group test is a very good method of ordering a number of children in respect of intelligence, with sufficient accuracy.

(2) The group test is not so reliable as the

individual test for assigning the exact mental level of a particular child. Though *in general* the agreement between the two tests is close, in *particular* cases the group test may rate a child too low or too high.

The factor which leads to the second conclusion seems to be that the group tests at present in use are not so comprehensive as the individual test in the appeal they make to the intelligence. The mental capacity of each child tested in the experiment here described was analysed as far as possible under thirteen heads,[1] all of which appear to be tested specifically by the S.B. scale, but some of which are not tested specifically by the group scale. In the cases of discrepancy scrutiny of the analysis revealed, almost invariably, that the child had

[1] These were: (1) Readiness and ability in applying knowledge; (2) Power to discriminate essentials; (3) Richness and logical integrity of the associative processes; (4) Ability to control and concentrate attention; (5) Power of comprehension; (6) Ability to hold in mind the conditions of a problem; (7) Ingenuity and practical judgment; (8) Steadiness of purpose; (9) Power of forming abstract ideas; (10) Power of generalization; (11) Critical ability; (12) Ability to manipulate imagery (especially visual imagery); (13) Development of social consciousness.

shown up well (or badly, as the case might be) in the individual test in certain directions which were not specifically covered by the group test. The remedy lies, of course, in making our group scales so comprehensive that we may be assured of their equal reliability with the individual scale, not only in general, but also in particular cases.[1]

II. The Direct Method. In Chapter III an experiment was described in which some 500 children were tested by the Terman group intelligence scale, and the method of deriving the children's mental ages and I.Q.s from their scores in the test was explained.

Now these children had also been tested by an ordinary written examination in arithmetic and English, so that data were available for comparing the results of the intelligence test with those of the ordinary educational test.

Before the comparison could be effected it was necessary to express the results of the written examination in terms similar to those in which the results of the intelligence test were expressed. This was done by means of a procedure exactly

[1] *Cf.* footnote at end of Chapter III.

analogous to that described in Chapter III. That is, the age norms of performance in the written examination (taking the arithmetic and English together) were established by finding the average scores of the children in the age groups $10\frac{1}{4}$, $10\frac{3}{4}$, $11\frac{1}{4}$, $11\frac{3}{4}$, $12\frac{1}{4}$, $12\frac{3}{4}$ years. A curve of performance was then obtained, and by comparing the score of a particular child with this curve it was possible to assign to him a ' mental age ' indicative not simply of native intelligence, but of educational attainment. Dividing this quantity by the child's actual age, and expressing the result as a percentage, a figure was obtained which might be fairly taken as an index of the child's educability, and was therefore termed the ' educability quotient ' (E.Q.).[1]

At this point it should be noted that the papers in the written examination were corrected in the ordinary way by a single examiner. A subjective factor was thus involved of a kind which did not enter into the intelligence test, the scoring of which was purely mechanical and independent of any particular examiner.

[1] *Cf.* Chapter II, p. 29.

It was therefore to be expected that the results of the written examination would not exhibit the same regularity as those of the intelligence test. This was in fact the case ; nevertheless the results were sufficiently regular to render the fitting of a curve of performance to the data a matter of little doubt or difficulty.

	Percentage of Children	
	Boys	Girls
E.Q. below I.Q. . . .	24.2	16.3
E.Q. and I.Q. corresponding .	55.7	56.1
E.Q. above I.Q. . . .	20.1	27.6

Now it is reasonable to suppose that if a child is educated in sufficiently favourable circumstances his attainment should keep pace with his intelligence, so that, *provided our intelligence tests are reliable*, a close correspondence should appear between the I.Q.s and the E.Q.s.

The results of this experiment are given in the above table (I.Q. and E.Q. are said to correspond when they are within five points of one another).

These figures speak for themselves, and it is

evident that the agreement between the results of the two kinds of test is strikingly close in a very considerable proportion of cases. Complete agreement could not in any case be expected, for many causes may militate against close agreement between the I.Q. and E.Q. of a particular child. Past or present illness, absence from school, late entry, bad teaching, laziness, will all tend to cause the E.Q. to drop below the I.Q. ; whereas a temperament more industrious than the average, good teaching, and similar causes may make the E.Q. rise above the I.Q., it being remembered that these two indexes are obtained by comparison with the *average*. In this experiment, of course, the average was not that of a sample of the whole population, but of this particular group of children, which (as explained in Chapter III) was somewhat above the average of the whole population. In obtaining the I.Q. it was of course necessary to use the *first* formula given in Chapter III.

The distributions were given by the following figures :

E.Q.	No. of Boys	No. of Girls	Total
Under 80 . . .	9	25	34
81–90 . . .	28	70	98
91–100 . . .	49	87	136
101–110 . . .	55	67	122
111–120 . . .	19	30	49
121–130 . . .	9	12	21
131 and over . .	6	6	12
	175	297	472

I.Q.	No. of Boys	No. of Girls	Total
Under 80 . . .	19	35	54
81–90 . . .	27	73	100
91–100 . . .	41	85	126
101–110 . . .	40	57	97
111–120 . . .	24	28	52
121–130 . . .	15	17	32
131 and over . .	8	2	10
	174	297	471

These figures give a median of just under 100 for both I.Q. and E.Q., but the semi-inter-quartile range shows a slightly greater scatter for the former than for the latter, a result

similar to that found in the experiment described in Chapter II. It probably points to a certain insufficiency in the elasticity of promotion so far as the brighter children are concerned.

Summing up, it may be said that the results thus obtained, both by the indirect and by the direct methods, afford good grounds for regarding tests of the group scale type as useful and sufficiently accurate means of estimating educability, while this reliability is likely to be increased when our tests are so devised as to probe the child's intelligence from as many directions as possible.

CHAPTER VI

CONCLUSION

WE may now consider, very briefly, the conclusions to which we are led by the results of the experiments which have been described. In the first place, it will probably be admitted that these experiments go some way toward establishing the reliability of intelligence tests in carrying out the special purpose for which they were designed—the separation and grading of children according to degrees of educability. Indeed, the general correspondence between the children's performances in the tests and the levels of educational attainment which they had shown themselves capable of reaching was really very striking indeed. It is therefore evident that intelligence tests should no longer be regarded with suspicion as a mere 'stunt,' but should be recognized as a valuable instrument capable of affording great aid in the advancement of the

theory and the practice of educational science. The task of the future is to perfect them.

Secondly, it must not be forgotten that intelligence is not the *only* factor necessary for success in school and in after life, though it is the most important. Unless he possesses a certain minimum of intelligence the child will not be capable of profiting by advanced education; but even if he exceeds this minimum other factors are necessary to his successful progress. Not only must he reach a certain standard of physical fitness, but his temperament and character must be such that he is impelled to use his intelligence effectively. It is just here that the real value of the ordinary educational tests of attainment becomes apparent. For we want to know not only whether a child is intelligent, but also whether he is capable of making good use of his intelligence. Intelligence tests are therefore most valuable when their results are scrutinized in comparison with the results of educational tests together with records of school careers. But if our educational tests are to attain their full value

an attempt must be made to standardize them as effectively as intelligence tests are standardized, and also to interpret their results, in the case of each child, by an appropriate age allowance carried out by some such method as that described in Chapter V. Unless dealt with in this way educational tests lose much of their significance. All cases in which a marked discrepancy is found between the results of the intelligence and of the attainment tests respectively should be made the subject of further study.

Finally, a word may be said as to the administration of the tests. While group testing can be carried out by any reasonably competent person who is willing to take the trouble to follow exactly the prescribed procedure (which is becoming more and more simple as the tests are improved), it is probable that the administration of the individual tests should be left to persons having some expert knowledge of the subject. In any case detailed interpretation of results is a matter for the psychologist. For this reason, as well as on account of the time

factor, practical testing in the future will inevitably (and quite rightly) be carried out mainly by the group method. The individual scale will remain as an ultimate standard of reference, to which recourse will be had for purposes of comparison, and for dealing with especially interesting and difficult cases, whether supernormal, subnormal, or abnormal.

INDEX

Administration of tests, 91
Age, mental, 17 f., 27, 40 ff., 56 ff.; critical, 56, 58, 63, 72; terminal, 56, 58, 72; effective, 63 f., 72; effective mental, 61 f., 72; effective terminal, 63, 72
Age performance, in arithmetic, 27, 83 f.; in mental tests, 41, 43 ff., 71; in composition, 28 ff., 71
American tests, 26, 36
Aptitude, special, 14
Arithmetic, 25, 36 f.; quotient, 27 ff.
Attainment, 16, 34 f., 90
Attention, 13

Ballard, P. B., 27 f.
Binet, Alfred, 24
Bravais, 33
British Journal of Psychology, 5
Brown, W., 21 *n.*
Burt, Cyril, 68

Character, 90
Coaching, for mental tests, 11 f.
Composition, 25, 27 f., 36 ff.; quotient, 29 ff.
Correlation, 20 f., 29 ff., 81
Critical age, 56, 63, 72
Criticisms of mental tests, 9 ff.
Curve of performance, 43 ff., 59 ff., 64 ff.

Defectiveness, mental, 32 f.
Discrepancy between results of different tests, 31, 53 f.
Distribution, 34; of intelligence quotients, 74 ff., 87; of educability quotients, 87; normal, 43 *n.*

Educability, 13, 16, 24 ff., 29 ff., 79 ff., 81, 84 ff., 89
Effective age, 63 f., 72
Effective mental age, 61 f., 72
Environment, 16
Error, probable, 20
Examinations, 83 ff., 90
Experience, 15

Factor, general intelligence, 13 f.; subjective, 17, 84 f.
Formula for mental age, 46 ff.

General intelligence factor, 13 f.
Growth of intelligence, 14 f., 57 ff.

Heterogeneity of material, 30 *n.*

Imagery, visual, 54
Intelligence, definition of, 12 f.; growth of, 14 f., 57 ff.; quotient, 17 f., 26 ff., 56 ff.; aspects of, 42, 82
Interest, 14
Interquartile range, 19

Knowledge, 15

Mean square deviation, 18 f.
Median, 19
Mental defectiveness, 32 f.

Normal distribution, 43 *n.*
Norms, 12, 44, 84

Objective standard, 17, 22
Ogive, 44 *n.*, 65 *n.*
Otis group test, 59 ff.

Pearson, K., 29, 33
Percentile rank, 18, 68, 74 ff.
Performance, curve of, 43 ff., 59 ff., 64 ff.
Personal equation, 17, 84 f.
Presupposition of mental tests, 17
Promotion, 34, 88.

Quartile, 19

Rank, percentile, 18, 68, 74 ff.
Records, school, 90

Scatter, 19, 34
Score in group test, 40 ff., 59 ff., 71; terminal, 63, 72
Semi-interquartile range, 19
Simplex Group Intelligence Scale, 55 *n.*
Spearman, C., 29
Standard deviation, 18 f.
Standardization, 37, 91
Stanford Revision of Binet tests, 24 ff., 43, 47, 66 f., 76 ff., 79, 80
Subjective factor, 17, 22

Teachers, estimates made by, 22
Temperament, 90
Terman, L. M., 26 *n.*, 66, 77 f.
Terman group test, 43, 47 ff., 71
Terminal age, 56, 58, 63, 72; score, 63, 72
Thomson, Godfrey, 21 *n.*

A LIST OF STANDARD INTELLIGENCE TESTS

TERMAN GROUP TEST OF MENTAL ABILITY

By Dr Lewis M. Terman.

For use in the upper standards of elementary schools and in secondary schools. Test can be completed in 35 minutes.

Examination Form A: 25 Examination Forms, one Key Form, one Manual of Directions, **6s.** 100 Examination Forms, one Key Form, one Manual of Directions, **21s.**

Examination Form B: Prices, etc., as for Form A.

Specimen Set, containing (1) one copy of Examination Forms A and B; (2) Keys for Forms A and B; (3) Manual of Directions. **2s.** net, in envelope.

OTIS GROUP INTELLIGENCE SCALE

By Arthur S. Otis.

An absolute point scale to measure in groups the mental ability of children throughout the elementary school and upward.

Primary Examination Forms A and B: Eight tests and a page for recording the performance of the examinee. Adapted for pupils from 6 to 9 years of age. 25 copies of either Form and one Class Record, **5s.** 100 Forms and one Class Record, **17s. 6d.**

Advanced Examination Forms A and B: Ten tests and a page for recording the performance of the examinee. Adapted for pupils of 10 years and upward. 25 copies of either Form and one Class Record, **5s.** 100 Forms and one Class Record, **17s. 6d.**

Manual of Directions: 72 pages. **1s. 6d.** net.

Examiner's Key for Advanced Examination: 20 pages. **1s.** net.

Specimen Set, containing (1) one copy of each of the Primary Forms and one each of the Advanced; (2) Manual of Directions; (3) Examiner's Key; (4) Class Record. **3s. 6d.** net, in envelope.

NATIONAL INTELLIGENCE TESTS

By Professors Haggerty, Terman, Thorndike, Whipple, and Yerkes.

An adaptation for English school purposes of the group intelligence tests used in the examination of recruits in the U.S. Army.

Scale A, Form 2: 25 Examination Forms and two Key Forms, **6s.** 100 Examination Forms and two Key Forms, **21s.**

Scale B, Form 2: 25 Examination Forms and one Key Form, **6s.** 100 Examination Forms and one Key Form, **21s.**

Manual of Directions: 1s. net.

Specimen Set, containing (1) Examination Forms A and B; (2) Keys for Forms A and B; (3) Manual of Directions. **2s. 6d.** net, in envelope.

THE SIMPLEX GROUP INTELLIGENCE SCALE

By C. A. RICHARDSON, M.A., Author of *Methods and Experiments in Mental Tests.*

No elaborate directions or timing are required in the use of this comprehensive Group Intelligence Scale, which is suitable for all ages above nine years.

Test Booklet: 25 copies, with Examiner's Report, **6s.** 100 copies, with 4 Examiner's Reports, **21s.**

Manual of Directions: 1s. net.

Specimen Set, containing (1) Test Booklet; (2) Manual of Directions; (3) Examiner's Stencil Key. **3s.** net, in envelope.

THE NORTHUMBERLAND MENTAL TESTS

By Professor GODFREY H. THOMSON, Armstrong College, Newcastle-on-Tyne.

A simple and very effective series.

Examination Form I: 25 copies, **5s.** 100 copies, **17s. 6d.**

Examination Form II: 25 copies, **5s.** 100 copies, **17s. 6d.**

Teacher's Manual: 1s. net.

Specimen Sets I and II, each containing (1) Examination Form, with Practice Test; (2) Teacher's Manual. **1s. 6d.** net, in envelope.

COURTIS STANDARD PRACTICE TESTS IN ARITHMETIC

By S. A. COURTIS, Supervisor of Educational Research in the Detroit Public Schools.

A series of forty-seven graded lessons on cards in two forms—A and B—including five research tests and eight cards for special study, covering the arithmetical operations with whole numbers for the elementary standards.

144 Practice Cards, sufficient for twelve pupils; in strong envelope, **7s. 6d. Student's Record and Practice Pad: 1s. Teacher's Manual: 1s. 6d.** net. **Lesson Cards, 1 to 48, Form A: 2s. 6d. Lesson Cards, 1 to 48, Form B: 2s. 6d.**

Specimen Set, containing (1) one set of Lesson Cards, 1 to 48, Form A; (2) one set of Lesson Cards, 1 to 48, Form B; (3) Student's Record and Practice Pad; (4) Teacher's Manual. **5s.** net, in envelope.

Full particulars of the above Tests and a list of New and Recent Books for Teachers will be sent on request.

GEORGE G. HARRAP & COMPANY LTD.
2 & 3 PORTSMOUTH STREET KINGSWAY W.C.
CALCUTTA AND SYDNEY

Date Due

OCT 11 337
OCT 18 "
ANov27 51 Hardy
Dec 12 '51 P Phillips
APR 27 1954 Campbell

22 JUL '58

JUL 31 1959

MAY 6 1960

R JUL 20 62

R AUG 16 62

11/3/65 101B
Azzara

MAR 22 REC'D

Property of Princeton
U. S. A.

CALL NO. BF 431 .R5
AUTHOR Richardson.
TITLE Methods and experimental tests.
Library of THE UNIVERSITY OF

DATE DUE | READER'S No
OCT 11 | 332
OCT 18 | "
JAN 27 | "
FEB 12 | "

UNIVERSITY OF B.C. LIBRARY
3 9424 02944 0580